ディリクレ過程事前分布を用いたベイズ推定

宮　脇　幸　治

I　はじめに

本論文ではディリクレ過程を事前分布に用いるベイズ推定について説明すると共に、その応用例としてサンプルセレクションモデルのベイズ推定について説明する。ベイズ推定とは、データを所与として未知パラメータの分布についての推測を行う統計的な方法であり、近年の計算機の発達と計算機集約的なマルコフ連鎖モンテカルロ法の普及に伴って、様々な統計的モデルの推測に用いられるようになってきた。ベイズ統計学に関する入門書としては、例えば松原（2008）及び安道（2010）、またファイナンスへの応用を念頭に置いたものとしては、例えば中妻（2013）が挙げられる。計量経済学への応用に関しては、和合（2005）が詳しい。

ベイズ推定について、簡単な例で見ておく。同じベルヌーイ分布に独立に従う N 個のデータ $\boldsymbol{X}=\{X_1, \ldots, X_N\}$ が得られているとし、$Y=\sum_{i=1}^{N} X_i$ とする。ベルヌーイ分布のパラメータ（$X_i=1$ となる確率、成功確率）を p とすれば、尤度関数は

$$f(\boldsymbol{X}|p)=\frac{\Gamma(N+1)}{\Gamma(Y+1)\Gamma(N-Y+1)}p^{Y}(1-p)^{N-Y}, \tag{1}$$

である。ただし $\Gamma(s)$ はガンマ関数。このパラメータ p に対して、ベータ分布 $Be(\alpha, \beta)$ を事前分布として仮定する。ここで $\alpha>0, \beta>0$。このとき p の事前密度は

$$\pi(p)=\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}p^{\alpha-1}(1-p)^{\beta-1}, \tag{2}$$

である。事後密度は尤度関数と事前密度の積に比例することを用いれば、事後密度は

$$\pi(p\,|\,\boldsymbol{X})\propto p^{\alpha+Y-1}(1-p)^{\beta+N-Y-1}, \tag{3}$$

となるので、事後分布は $Be(\alpha+Y, \beta+N-Y)$ であることが分かる。事後分布を解析的に分析できない場合は、マルコフ連鎖モンテカルロ法などの手法で事後分布からの乱数を発生させ、統計的推測を行う。事後分布が得られた後は、点推定（事後平均など)、区間推定（95%信用区間など）を行ったり、新しいデータ z に対する予測分布 $f(z\,|\,\boldsymbol{X})=\int f(z\,|\,p)\pi(p\,|\,\boldsymbol{X})dp$ を求めたりすることで、分析を行う。今の例において、新しく1個のデータ z が得られたと考えたとき、それに対する予測分布は成功確率 $(\alpha+Y)/(\alpha+\beta+N)$ のベルヌーイ分布となる。

簡単な例で見たように、ベイズ推定では統計的モデルがパラメトリックなものを用いることが特徴である。しかし近年ノンパラメトリックな密度推定やセミパラメトリックなモデルの推定においても、ベイズ推定を行うことが提案されている。そこで主に用いられている確率過程はディリクレ過程と呼ばれるものであり、それは確率分布に対する分布である。以下の節で詳しく見ていくが、ディリクレ過程を事前分布として用いることで、ノンパラメトリックな密度推定やセミパラメトリックなモデルの推定を行うことが出来る。

本論文では、まずⅡ節でディリクレ過程の定義を説明し、続くⅢ節でディリクレ過程を事前分布とした場合のベイズ推定について簡単に紹介する。Ⅳ節では、Ishwaran and Zarepour（2000）によって提案されたディリクレ過程の近似を紹介し、実際に回帰モデルの推定方法について説明する。サンプルセレクションモデルへの応用についてはⅤ節で見る。Ⅵ節で本論文をまとめる。

II　ディリクレ過程とは

ディリクレ過程とは、Ferguson（1973）によって提案された確率過程で、確率分布に対する分布である。以下にその定義を与える。ここでの定義は古澄（2005）に従っている。

定義 1． 任意の可測空間（$X, \mathcal{B}$）を考える。また正のスカラー α と（$X, \mathcal{B}$）上で定義された確率分布 G_0 を考える。G が α と G_0 をパラメータとするディリクレ過程であるとは、X の任意の分割 $B_1, \ldots, B_m$ に対して、$\{P(B_1), \ldots, P(B_m)\}$ がパラメータ $\{\alpha G_0(B_1), \ldots, \alpha G_0(B_m)\}$ のディリクレ分布に従うことであり、$G \sim DP(\alpha G_0)$ と表記する。

ここで分割 $B_1, \ldots, B_m$ とは、$B_i \in \mathcal{B}$, $B_i \cap B_j = \emptyset$, $\cup_{i=1}^{m} B_i = \mathcal{B}$ を満たしているものである。また、ディリクレ分布の密度関数は

$$\frac{\Gamma(\alpha)}{\prod_{i=1}^{m}\Gamma\{\alpha G_0(B_i)\}}\prod_{i=1}^{m} G(B_i)^{\alpha G_0(B_i)-1}, \tag{4}$$

で与えられる。特に $m=2$ とすれば、任意の $B \in \mathcal{B}$ に対して、パラメータ（$\alpha G_0(B), \alpha\{1-G_0(B)\}$）のディリクレ分布を考えることになるが、これはベータ分布 $Be(\alpha G_0(B), \alpha\{1-G_0(B)\}$）である。ディリクレ分布のこの性質を用いれば、ディリクレ過程の期待値と分散を求めることが出来る。つまり、ディリクレ過程に従う G の期待値と分散はそれぞれ

$$E\{G(B)\} = G_0(B), \quad Var\{G(B)\} = \frac{G_0(B)\{1-G_0(B)\}}{\alpha+1}, \tag{5}$$

となる[1)]。よってディリクレ過程のパラメータ α は G の分散に影響するパラ

---

1)　ベータ分布 $Be(\alpha, \beta)$ に従う確率変数 X の期待値と分散はそれぞれ $E(X) = \dfrac{\alpha}{\alpha+\beta}$, $Var(X) = \dfrac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$ である。

メータであり、G_0 は G の期待値を決めるパラメータであることが分かる。後者は基底測度と呼ばれる。その他のディリクレ過程に関する性質については、例えば Ghosal (2010) を参照されたい。

Ⅲ　ディリクレ過程を事前分布に用いたベイズ推定

以下の回帰モデルを考える。つまり N 個の観測値に対して

$$y_i \mid \boldsymbol{x}_i \sim N(\boldsymbol{x}_i' \boldsymbol{\beta}, \sigma^2), \quad i=1, \ldots, N. \tag{6}$$

下付きの i は i 番目の観測値であることを表す。また y_i は被説明変数、$\boldsymbol{x}_i$ と $\boldsymbol{\beta}$ は説明変数ベクトルと係数ベクトルであり、y_i は平均 $\boldsymbol{x}_i'\boldsymbol{\beta}$、分散 σ^2 の条件付き正規分布に独立に従う。つまり σ^2 は条件付き分布の分散を表すパラメータとなっている。通常のベイズ推定では、パラメータの事前分布として、例えば $\boldsymbol{\beta}$ には多変量正規分布、σ^2 には逆ガンマ分布を仮定することで、事後分布を導出し、統計的推測を行う。

このモデルは y_i の条件付き分布に正規分布を仮定しているが、この仮定をディリクレ過程を用いることで、以下のように緩める。つまり

$$\begin{aligned} y_i \mid \boldsymbol{x}_i &\sim N(\boldsymbol{x}_i' \boldsymbol{\beta}_i, \sigma_i^2), \ i=1, \ldots, N, \\ \phi_i \equiv (\boldsymbol{\beta}_i', \sigma_i^2)' \mid G &\sim G, \\ G &\sim DP(\alpha G_0). \end{aligned} \tag{7}$$

このモデルに対するベイズ推定を行うには、マルコフ連鎖モンテカルロ法を用いて、事後分布からの乱数を発生させる必要がある。その際ディリクレ過程の以下の性質を用いる。つまり

$$\begin{aligned} \phi_i \mid G &\sim G, \ i=1, \ldots, N, \\ G &\sim DP(\alpha G_0), \end{aligned} \tag{8}$$

であるとき、G を周辺化すれば

$$\begin{aligned} \phi_1 &\sim G_0, \\ \phi_i \mid \phi_1, \ldots, \phi_{i-1} &\sim \frac{\alpha G_0 + \sum_{j=1}^{i-1} \delta_{\phi_j}}{\alpha + N + 1}, \ i=2, \ldots, N, \end{aligned} \tag{9}$$

となる。ただし δ_x は点 x で 1 、それ以外では 0 となる関数である。この表現より、$\phi_i(i\geq 2)$ は、それ以前の $\phi_1, \ldots, \phi_{i-1}$ を所与として、確率 $1/(\alpha+N-1)$ で ϕ_j、確率 $\alpha/(\alpha+N-1)$ で基底測度 G_0 から新しく発生された値となることが分かる。このような表現は、一般化 Pólya の壺の枠組みと呼ばれている（Blackwell and MacQueen（1973））。モデル（7）に対して、この枠組みを用いたベイズ推定の実際については、古澄（2005）を参照されたい。

Ⅳ　Ishwaran and Zarepour（2000）による近似

ディリクレ過程に従う分布 G は、以下のように表現することも出来る。つまり、$G\sim DP(\alpha, G_0)$ であるとき

$$
\begin{aligned}
&G=\sum_{k=1}^{\infty} p_k \delta_{\phi_k^\dagger},\\
&p_1=V_1,\ p_k=V_k\prod_{l=1}^{k-1}(1-V_l),\ (k\geq 2), \qquad (10)\\
&V_k\sim Be(1, \alpha),\ \phi_k^\dagger\sim G_0,
\end{aligned}
$$

と表現できる。ただし全ての V_k と $\phi_k^\dagger$ は独立である。これは Sethuraman（1994）によって提案された表現方法であり、stick-breaking 表現と呼ばれる。

もしパラメータ $\{\phi_i\}_{i=1}^N$ が独立に G に従っていれば、これらは $\{\phi_k^\dagger\}_k$ のいずれかに一致する。このとき、ϕ_i と $\phi_j(i\neq j)$ が同じ $\phi_k^\dagger$ となることを許しているため、パラメータ $\{\phi_i\}_{i=1}^N$ は、いくつかの $\phi_k^\dagger$ によってグループ分けされることが分かる。つまりディリクレ過程は $\{\phi_i\}_{i=1}^N$ を未知の数のグループに分ける性質を持つことが分かる。

Stick-breaking 表現は無限和を含む形で得られているが、Ishwaran and Zarepour（2000）では、この無限和を有限和で近似する表現を提案している。つまり

$$
G_K=\sum_{k=1}^{K} p_k \delta_{\phi_k^\dagger}, \qquad (11)
$$

とし、その他は同じとする。この近似は概切断（almost sure truncation）と呼ばれる。

Ishwaran and Zarepour（2000）では近似の程度についても議論されており、ディリクレ過程のパラメータの値にもよるが、K が十分大きいとき、この有限和による近似表現は元のディリクレ過程に十分近いことが数値例によって示されている。

この近似表現を用いて、先のモデル（7）のベイズ推定を考える。そのために、新しい潜在変数 s_i^* を導入する。この変数は 1 から K の間の値を取る離散変数で、ϕ_i がどの $\{\phi_k^\dagger\}_k$ に一致するかを示す変数である。つまり $\phi_i = \phi_k^\dagger$ のとき、$s_i^* = k$ である。このとき、G に代えて近似表現 G_K を用いたモデルは

$$
\begin{aligned}
y_i \mid \boldsymbol{x}_i &\sim N\left\{\boldsymbol{x}_i'\boldsymbol{\beta}_{s_i^*}^\dagger, \left(\sigma_{s_i^*}^\dagger\right)^2\right\}, \\
\Pr(s_i^* = s \mid \boldsymbol{p}) &\propto p_s, \\
f(\boldsymbol{p}) &\propto \alpha^{K-1} p_K^{\alpha-1}(1-p_1)^{-1}\cdots(1-p_1-\cdots p_{K-2})^{-1}, \\
\phi_k^\dagger \equiv ((\boldsymbol{\beta}_k^\dagger)', (\sigma_k^\dagger)^2)' &\sim G_0,
\end{aligned} \tag{12}
$$

となる。ただし $\boldsymbol{p} \equiv (p_1, \ldots, p_K)'$ である。ベクトル $\boldsymbol{p}$ の同時分布は一般化ディリクレ分布と呼ばれ、$2(K-1)$ 個のパラメータを持ち、$GD(1, \alpha, \ldots, 1, \alpha)$ と表記される[2)]。この近似表現の導出に関する詳細は Ishwaran and Zarepour（2000）や Ishwaran and James（2001）を参照されたい。

以下の事前分布を仮定すれば、モデル（12）のベイズ推定はマルコフ連鎖モンテカルロ法によって発生された乱数を用いて行えば良い。事前分布として

$$
dG_0 = N(\boldsymbol{\mu}, \boldsymbol{\Sigma})IG\left(\frac{n_0}{2}, \frac{\omega}{2}\right), \tag{13}
$$

$$
\boldsymbol{\mu} \sim N(\boldsymbol{m}_0, \boldsymbol{S}_0),\ \boldsymbol{\Sigma} \sim IW(r_0, \boldsymbol{R}_0),\ \alpha \sim G(a_0, b_0),\ \omega \sim G\left(\frac{c_0}{2}, \frac{d_0}{2}\right), \tag{14}
$$

を仮定する[3)]。ここで $\boldsymbol{\beta}_k^\dagger$ と $(\sigma_k^\dagger)^2$ の基底測度には、正規分布と逆ガンマ分布

2）一般化ディリクレ分布については、例えば Connor and Mosimann（1969）を参照。
3）正の確率変数 X がガンマ分布に従うとき、$X \sim G(\alpha, \beta)$ と表し、その密度関数は

を仮定している。また集合 $\mathcal{N}_k=\{i|s_i^*=k\}$ とその要素数 N_k を $k=1, \dots, K$ に対して定義する。このとき、マルコフ連鎖モンテカルロ法による乱数発生アルゴリズムは、以下のステップで与えられる。

Step 1. $\{\boldsymbol{\beta}_k^\dagger, (\sigma_k^\dagger)^2\}_{k=1}^K$, $\{s_i^*\}_{i=1}^N$, $\boldsymbol{p}$, $\boldsymbol{\mu}$, $\boldsymbol{\Sigma}$, ω, α に適当な初期値を与える。

Step 2. $\boldsymbol{\beta}_k^\dagger(k=1, \dots, K)$ を以下の条件付き分布から発生させる。

$$\boldsymbol{\beta}_k^\dagger|(\sigma_k^\dagger)^2, \{s_i^*\}_{i=1}^N, \boldsymbol{\mu}, \boldsymbol{\Sigma}\sim N(\boldsymbol{m}_{k,1}, \boldsymbol{S}_{k,1}), \tag{17}$$

ただし

$$\boldsymbol{S}_{k,1}^{-1}=\boldsymbol{\Sigma}^{-1}+(\sigma_k^\dagger)^{-2}\sum_{i\in\mathcal{N}_k}\boldsymbol{x}_i\boldsymbol{x}_i', \boldsymbol{m}_{k,1}=\boldsymbol{S}_{k,1}\left(\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}+(\sigma_k^\dagger)^{-2}\sum_{i\in\mathcal{N}_k}\boldsymbol{x}_i y_i\right). \tag{18}$$

Step 3. $(\sigma_k^\dagger)^2$（$k=1, \dots, K$）を以下の条件付き分布から発生させる。

$$(\sigma_k^\dagger)^2|\boldsymbol{\beta}_k^\dagger, \{s_i^*\}_{i=1}^N, \omega\sim IG\left(\frac{n_0+N_k}{2}, \frac{\omega+\sum_{i\in\mathcal{N}_k}(y_i-\boldsymbol{x}_i'\boldsymbol{\beta}_k^\dagger)^2}{2}\right), \tag{19}$$

Step 4. s_i^*（$i=1, \dots, N$）を以下の条件付き分布から発生させる。

$$\Pr(s_i^*=s|\{\boldsymbol{\beta}_k^\dagger, (\sigma_k^\dagger)^2\}_{k=1}^K, \boldsymbol{p})\propto\frac{p_s}{\sigma_s^\dagger}\exp\left\{-\frac{1}{2(\sigma_s^\dagger)^2}(y_i-\boldsymbol{x}_i'\boldsymbol{\beta}_s^\dagger)^2\right\}. \tag{20}$$

Step 5. $\boldsymbol{p}$ を以下の条件付き分布から発生させる。

$$\boldsymbol{p}|\{s_i^*\}_{i=1}^N, \alpha\sim GD(1+N_1, \alpha+M_1, \dots, 1+N_{K-1}, 1+M_{K-1}). \tag{21}$$

$$\frac{\beta^\alpha}{\Gamma(\alpha)}x^{\alpha-1}\exp(-\beta x), \tag{15}$$

で与えられる。ただし $\alpha>0, \beta>0$。このとき期待値と分散はそれぞれ α/β 及び α/β^2 である。また逆数 X^{-1} は逆ガンマ分布 $IG(\alpha, \beta)$ に従うことが知られており、逆ガンマ分布の期待値と分散はそれぞれ $\beta/(\alpha-1)$ 及び $\beta^2/\{(\alpha-1)^2(\alpha-2)\}$ である（ただし $\alpha>2$）。

対称で正値定符号な $p\times p$ 行列 $\boldsymbol{V}$ が逆ウィッシャート分布に従うとき、$IW(m, \boldsymbol{\Psi})$ と表し、その密度関数は

$$\begin{aligned} &c|\boldsymbol{\Psi}|^{(m-p-1)/2}|\boldsymbol{V}|^{-m/2}\exp\left\{-\frac{1}{2}\mathrm{tr}(\boldsymbol{\Psi}\boldsymbol{V}^{-1})\right\}, \\ &c^{-1}=2^{(m-p-1)p/2}\pi^{p(p-1)/4}\prod_{i=1}^p\Gamma\left(\frac{m-p-i}{2}\right), \end{aligned} \tag{16}$$

で与えられる。ただし $m>2p, \boldsymbol{\Psi}>0$。期待値は $m-2p-2>0$ のとき存在し $\boldsymbol{\Psi}/(m-2p-2)$ である。

ただし $M_k=\sum_{h=k+1}^{K}N_h$。この条件付き分布に従う乱数を得るには、以下の逐次的な手順を用いる。

1. V_1 を $Be(1+N_1, \alpha+M_1)$ から発生させ、$p_1=V_1$ とする。

2. V_2 を $Be(1+N_2, \alpha+M_2)$ から発生させ、$p_2=(1-V_1)V_2$ とする。

…

K-1. V_{K-1} を $Be(1+N_{K-1}, \alpha+M_{K-1})$ から発生させ、$p_{K-1}=(1-V_1)\ldots(1-V_{K-2})V_{K-1}$ とする。

K. $V_K=1$ とし、$p_K=(1-V_1)\ldots(1-V_{K-1})V_K$ とする。

Step 6. $\boldsymbol{\mu}$ を以下の条件付き分布から発生させる。

$$\boldsymbol{\mu}\,|\,\{\boldsymbol{\beta}_k^{\dagger}\}_{k=1}^{K}, \boldsymbol{\Sigma}\sim N(\boldsymbol{m}_1, \boldsymbol{S}_1), \tag{22}$$

ただし

$$\boldsymbol{S}_1^{-1}=\boldsymbol{S}_0^{-1}+K\cdot\boldsymbol{\Sigma}^{-1}, \boldsymbol{m}_1=\boldsymbol{S}_1\left(\boldsymbol{S}_0^{-1}\boldsymbol{m}_0+\boldsymbol{\Sigma}^{-1}\sum_{k=1}^{K}\boldsymbol{\beta}_k^{\dagger}\right). \tag{23}$$

Step 7. $\boldsymbol{\Sigma}$ を以下の条件付き分布から発生させる。

$$\boldsymbol{\Sigma}\,|\,\{\boldsymbol{\beta}_k^{\dagger}\}_{k=1}^{K}, \boldsymbol{\mu}\sim IW\left\{r_0+K, \boldsymbol{R}_0+\sum_{k=1}^{K}(\boldsymbol{\beta}_k^{\dagger}-\boldsymbol{\mu})(\boldsymbol{\beta}_k^{\dagger}-\boldsymbol{\mu})'\right\}. \tag{24}$$

Step 8. ω を以下の条件付き分布から発生させる。

$$\omega\,|\,\{(\sigma_k^{\dagger})^2\}_{k=1}^{K}\sim G\left\{\frac{c_0+K\cdot n_0}{2}, \frac{d_0+\sum_{k=1}^{K}(\sigma_k^{\dagger})^{-2}}{2}\right\}. \tag{25}$$

Step 9. α を以下の条件付き分布から発生させる。

$$\alpha\,|\,p_K\sim\begin{cases}G(a_0+K-1, b_0-\log p_k), & \text{if } p_K>0,\\ G(a_0, b_0), & \text{if } p_K=0.\end{cases} \tag{26}$$

Step 10. Step 2 に戻る。

V　サンプルセレクションモデルへの応用

本節では、ディリクレ過程事前分布を用いたサンプルセレクションモデルの推定手法について説明する。ベースとなるサンプルセレクションモデルは、Omori（2007）で用いられているものであり、$i=1, \ldots, N$ に対して

$$
\begin{aligned}
&y_i=\begin{cases} y_i^*, & \text{if } z_i^*>0,\\ \text{n.a.}, & \text{otherwise,}\end{cases}\\
&z_i^*=\boldsymbol{w}_i'\boldsymbol{\theta}+u_i,\\
&y_i^*=\boldsymbol{x}_i'\boldsymbol{\beta}+v_i.
\end{aligned}
\tag{27}
$$

被説明変数 y_i は潜在変数 z_i^* が正かどうかに依存しており、正であれば潜在変数 y_i^* が観測され、そうでなければ打ち切られる。これらの潜在変数 z_i^*, y_i^* を説明する変数ベクトルとして $\boldsymbol{w}_i, \boldsymbol{x}_i$ があり、$\boldsymbol{\theta}$ と $\boldsymbol{\beta}$ はそれぞれ対応する係数ベクトルとなっている。誤差項 (u_i, v_i) には二変量正規分布を仮定する。Amemiya（1984）の分類に従えば、第Ⅱ型の Tobit モデルと呼ばれることもある。

ディリクレ過程を用いて、潜在変数に関する統計的モデルを以下のように拡張する。つまり

$$
\begin{aligned}
&z_i^*=\boldsymbol{w}_i'\boldsymbol{\theta}_i+u_i,\\
&y_i^*=\boldsymbol{x}_i'\boldsymbol{\beta}_i+v_i,\\
&(u_i, v_i)'\sim N(\boldsymbol{0}, \boldsymbol{\Sigma}_i),\ \boldsymbol{\Sigma}_i\equiv\begin{pmatrix}1 & \gamma_i\\ \gamma_i & \phi_i+\gamma_i^2\end{pmatrix},\\
&(\boldsymbol{\theta}_i', \boldsymbol{\beta}_i', \gamma_i, \phi_i)'\,|\,G\sim G,\\
&G\sim DP(\alpha G_0).
\end{aligned}
\tag{28}
$$

拡張されたモデルでは、潜在変数における説明変数の係数と誤差項の分散共分散行列が観測値によって変動することとなり、その同時事前分布がディリクレ過程に従うと仮定されている。

拡張された統計的モデルに対して、前節で説明した stick-breaking 表現とその近似を導入すれば

$$
\begin{aligned}
&y_i=\begin{cases} y_i^*, & \text{if } z_i^*>0,\\ \text{n.a.}, & \text{otherwise,}\end{cases}\\
&z_i^*=\boldsymbol{w}_i'\boldsymbol{\theta}_{s_i^*}^{\dagger}+u_i,\\
&y_i^*=\boldsymbol{x}_i'\boldsymbol{\beta}_{s_i^*}^{\dagger}+v_i,
\end{aligned}
$$

$$(u_i, v_i)' \sim N\left(0, \boldsymbol{\Sigma}^{\dagger}_{s_i^*}\right),\ \boldsymbol{\Sigma}^{\dagger}_{s_i^*} \equiv \begin{pmatrix} 1 & \gamma^{\dagger}_{s_i^*} \\ \gamma^{\dagger}_{s_i^*} & \phi^{\dagger}_{s_i^*} + \left(\gamma^{\dagger}_{s_i^*}\right)^2 \end{pmatrix}, \tag{29}$$

$$\Pr(s_i^* = s \mid \boldsymbol{p}) \propto p_s,$$

$$\boldsymbol{p} \sim GD(1, \alpha, \ldots, 1, \alpha),$$

$$((\boldsymbol{\theta}^{\dagger}_k)', (\boldsymbol{\beta}^{\dagger}_k)', \gamma^{\dagger}_k, \phi^{\dagger}_k)' \sim G_0,$$

となる。

ベイズ推定を行うために、以下の事前分布を仮定する。

$$dG_0 = N(\boldsymbol{\mu}_\theta, \boldsymbol{\Sigma}_\theta) N(\boldsymbol{\mu}_\beta, \boldsymbol{\Sigma}_\beta) N(\mu_\gamma, \sigma_\gamma^2) IG\left(\frac{n_0}{2}, \frac{\omega}{2}\right), \tag{30}$$

$$\begin{aligned} &\boldsymbol{\mu}_\theta \sim N(\boldsymbol{m}_{\theta,0}, \boldsymbol{S}_{\theta,0}),\ \boldsymbol{\Sigma}_\theta \sim IW(r_{\theta,0}, R_{\theta,0}),\ \boldsymbol{\mu}_\beta \sim N(\boldsymbol{m}_{\beta,0}, \boldsymbol{S}_{\beta,0}), \\ &\quad \boldsymbol{\Sigma}_\beta \sim IW(r_{\beta,0}, \boldsymbol{R}_{\beta,0}), \end{aligned} \tag{31}$$

$$\begin{aligned} &\mu_\gamma \sim N(m_{\gamma,0}, s_{\gamma,0}^2),\ \sigma_\gamma^2 \sim IG\left(\frac{r_{\gamma,0}}{2}, \frac{R_{\gamma,0}}{2}\right),\ \alpha \sim G(a_0, b_0), \\ &\quad \omega \sim G\left(\frac{c_0}{2}, \frac{d_0}{2}\right). \end{aligned} \tag{32}$$

ここで、基底測度として $\boldsymbol{\theta}^{\dagger}_k$ と $\boldsymbol{\beta}^{\dagger}_k$ にはそれぞれ $N(\boldsymbol{\mu}_\theta, \boldsymbol{\Sigma}_\theta)$, $N(\boldsymbol{\mu}_\beta, \boldsymbol{\Sigma}_\beta)$ なる多変量正規分布、$\gamma^{\dagger}_k$ には（一変量）正規分布、$\phi^{\dagger}_k$ には逆ガンマ分布を仮定している。また集合 $\mathcal{NC}_k = \{i \mid y_i$ は観測されており $s_i^* = k\}$, $C_k = \{i \mid y_i$ は打ち切られており $s_i^* = k\}$ と定義する。このとき、以下のマルコフ連鎖モンテカルロ法のアルゴリズムを用いて、事後分布からの乱数を発生させ、統計的推測を行うことができる。

Step 1. $\{\boldsymbol{\theta}^{\dagger}_k, \boldsymbol{\beta}^{\dagger}_k, \gamma^{\dagger}_k, \phi^{\dagger}_k\}_{k=1}^K$, $\{s_i^*, z_i^*\}_{i=1}^N$, $\boldsymbol{p}$, $\boldsymbol{\mu}_\theta$, $\boldsymbol{\Sigma}_\theta$, $\boldsymbol{\mu}_\beta$, $\boldsymbol{\Sigma}_\beta$, μ_γ, σ_γ^2, ω, α に適当な初期値を与える。

Step 2. $(\boldsymbol{\theta}^{\dagger}_k, \boldsymbol{\beta}^{\dagger}_k)$ $(k=1, \ldots, K)$ を以下の条件付き分布から発生させる。

$$\begin{pmatrix} \boldsymbol{\theta}^{\dagger}_k \\ \boldsymbol{\beta}^{\dagger}_k \end{pmatrix} \Big| \{s_i^*, z_i^*\}_{i=1}^N, \gamma^{\dagger}_k, \phi^{\dagger}_k, \boldsymbol{\mu}_\theta, \boldsymbol{\Sigma}_\theta, \boldsymbol{\mu}_\beta, \boldsymbol{\Sigma}_\beta, \sim N(m_{k,1}, S_{k,1}), \tag{33}$$

ただし

$$\boldsymbol{S}_{k,1}^{-1}=\boldsymbol{S}_0^{-1}+\sum_{i\in \mathcal{N}C_k}\tilde{\boldsymbol{X}}_i'(\boldsymbol{\Sigma}_k^{\dagger})^{-1}\tilde{\boldsymbol{X}}_i+\sum_{i\in C_k}\tilde{\boldsymbol{X}}_i'\tilde{\boldsymbol{X}}_i, \tag{34}$$

$$\boldsymbol{m}_{k,1}=\boldsymbol{S}_{k,1}\left\{\boldsymbol{S}_0^{-1}\boldsymbol{m}_0+\sum_{i\in \mathcal{N}C_k}\tilde{\boldsymbol{X}}_i'(\boldsymbol{\Sigma}_k^{\dagger})^{-1}\tilde{\boldsymbol{y}}_i+\sum_{i\in C_k}\tilde{\boldsymbol{X}}_i'\tilde{\boldsymbol{y}}_i\right\}, \tag{35}$$

$$\boldsymbol{S}_0=\begin{pmatrix}\boldsymbol{\Sigma}_\theta & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{\Sigma}_\beta\end{pmatrix},\ \boldsymbol{m}_0=\begin{pmatrix}\boldsymbol{\mu}_\theta\\ \boldsymbol{\mu}_\beta\end{pmatrix},$$

$$[\tilde{\boldsymbol{y}}_i,\ \tilde{\boldsymbol{X}}_i]=\begin{cases}\left[\begin{pmatrix}z_i^*\\ y_i\end{pmatrix},\begin{pmatrix}\boldsymbol{w}_i' & \boldsymbol{0}'\\ \boldsymbol{0}' & \boldsymbol{x}_i'\end{pmatrix}\right], & \text{if } z_i^*>0,\\ \left[\begin{pmatrix}z_i^*\\ 0\end{pmatrix},\begin{pmatrix}\boldsymbol{w}_i' & \boldsymbol{0}'\\ \boldsymbol{0}' & \boldsymbol{0}'\end{pmatrix}\right], & \text{otherwise}\end{cases} \tag{36}$$

Step 3. $\gamma_k^{\dagger}$（$k=1, \ldots, K$）を以下の条件付き分布から発生させる。

$$\gamma_k^{\dagger}|\{s_i^*, z_i^*\}_{i=1}^N, \boldsymbol{\theta}_k^{\dagger}, \boldsymbol{\beta}_k^{\dagger}, \phi_k^{\dagger}, \mu_\gamma, \sigma_\gamma^2\sim N(\tilde{m}_{k,1}, \tilde{s}_{k,1}^2), \tag{37}$$

ただし

$$\tilde{s}_{k,1}^{-2}=\sigma_\gamma^{-2}+(\phi_k^{\dagger})^{-1}\sum_{i\in \mathcal{N}C_k}(z_i^*-\boldsymbol{w}_i'\boldsymbol{\theta}_k^{\dagger})^2, \tag{38}$$

$$\tilde{m}_{k,1}=\tilde{s}_{k,1}^2\left\{\sigma_\gamma^{-2}\mu_\gamma+(\phi_k^{\dagger})^{-1}\sum_{i\in \mathcal{N}C_k}(z_i^*-\boldsymbol{w}_i'\boldsymbol{\theta}_k^{\dagger})(y_i-\boldsymbol{x}_i'\boldsymbol{\beta}_k^{\dagger})\right\}. \tag{39}$$

Step 4. $\phi_k^{\dagger}$（$k=1, \ldots, K$）を以下の条件付き分布から発生させる。

$$\phi_k^{\dagger}|\{s_i^*, z_i^*\}_{i=1}^N, \boldsymbol{\theta}_k^{\dagger}, \boldsymbol{\beta}_k^{\dagger}, \gamma_k^{\dagger}, \omega\sim IG\left(\frac{n_0+NC_k}{2}, \frac{R_{k,1}}{2}\right), \tag{40}$$

ただし NC_k は $\mathcal{N}C_k$ の要素数、

$$\begin{aligned}R_{k,1}=\omega&+(\gamma_k^{\dagger})^2\sum_{i\in \mathcal{N}C_k}(z_i^*-\boldsymbol{w}_i'\boldsymbol{\theta}_k^{\dagger})^2-2\gamma_k^{\dagger}\sum_{i\in \mathcal{N}C_k}(z_i^*-\boldsymbol{w}_i'\boldsymbol{\theta}_k^{\dagger})(y_i-\boldsymbol{x}_i'\boldsymbol{\beta}_k^{\dagger})\\&+\sum_{i\in \mathcal{N}C_k}(y_i-\boldsymbol{x}_i'\boldsymbol{\beta}_k^{\dagger})^2.\end{aligned} \tag{41}$$

Step 5. (s_i^*, z_i^*)（$i=1, \ldots, N$）を以下の条件付き分布から発生させる。まず s_i^* を以下の 条件つき多項分布より発生させる。

$$\Pr(s_i^*=s\,|\,\boldsymbol{\theta}_s^{\dagger}, \boldsymbol{\beta}_s^{\dagger}, \gamma_s^{\dagger}, \phi_s^{\dagger}, p_s)\propto L_{is}^*p_s, \tag{42}$$

ただし $\Phi(z)$ は標準正規分布に従う確率変数 Z の累積分布関数、

$$L^*_{is}=\begin{cases}\{\phi^\dagger_s+(\gamma^\dagger_s)^2\}^{-1/2}\left\{1-\Phi\left(-\dfrac{m_{z,1}}{s_{z,1}}\right)\right\}\exp\left\{-\dfrac{1}{2}\dfrac{(y_i-\boldsymbol{x}_i'\boldsymbol{\beta}^\dagger_s)^2}{\phi^\dagger_s+(\gamma^\dagger_s)^2}\right\}, & \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{if } i\in\mathcal{NC}_s, & \\ \Phi(-\boldsymbol{w}_i'\boldsymbol{\theta}^\dagger_s) \qquad\qquad\qquad\qquad\qquad\qquad\quad \text{if } i\in C_s & \end{cases} \tag{43}$$

$$s^2_{z,1}=1-\frac{(\gamma^\dagger_s)^2}{\phi^\dagger_s+(\gamma^\dagger_s)^2},\ m_{z,1}=\boldsymbol{w}_i'\boldsymbol{\theta}^\dagger_s+\frac{\gamma^\dagger_s(y_i-\boldsymbol{x}_i'\boldsymbol{\beta}^\dagger_s)}{\phi^\dagger_s+(\gamma^\dagger_s)^2}. \tag{44}$$

続いて、得られた $s^*_i=s$ を所与として、z^*_i を以下の切断正規分布より発生させる。

$$z^*_i\,|\,s^*_i,\,\boldsymbol{\beta}^\dagger_s,\,\boldsymbol{\theta}^\dagger_s,\,\gamma^\dagger_s,\,\phi^\dagger_s\sim\begin{cases}N(m_{z,1},\,s^2_{z,1})I(z^*_i>0), & \text{if } i\in\mathcal{NC}_s,\\ N(\boldsymbol{w}_i'\boldsymbol{\theta}^\dagger_s,\,1)I(z^*_i\le 0), & \text{if } i\in C_s,\end{cases} \tag{45}$$

Step 6. $\boldsymbol{p}$ の発生はⅣ節と同じ。

Step 7. $\boldsymbol{\mu}_\theta$ を以下の条件付き分布から発生させる。

$$\boldsymbol{\mu}_\theta\,|\,\{\boldsymbol{\theta}^\dagger_k\}^K_{k=1},\,\boldsymbol{\Sigma}_\theta\sim N(\boldsymbol{m}_{\theta,1},\,\boldsymbol{S}_{\theta,1}), \tag{46}$$

ただし

$$\boldsymbol{S}^{-1}_{\theta,1}=\boldsymbol{S}^{-1}_{\theta,0}+K\cdot\boldsymbol{\Sigma}^{-1}_\theta,\ \boldsymbol{m}_{\theta,1}=\boldsymbol{S}_{\theta,1}\left(\boldsymbol{S}^{-1}_{\theta,0}\boldsymbol{m}_{\theta,0}+\boldsymbol{\Sigma}^{-1}_\theta\sum_{k=1}^{K}\theta^\dagger_k\right). \tag{47}$$

Step 8. $\boldsymbol{\Sigma}_\theta$ を以下の条件付き分布から発生させる。

$$\boldsymbol{\Sigma}_\theta\,|\,\{\boldsymbol{\theta}^\dagger_k\}^K_{k=1},\,\boldsymbol{\mu}_\theta\sim IW\left\{r_{\theta,0}+K,\,\boldsymbol{R}_{\theta,0}+\sum_{k=1}^{K}(\boldsymbol{\theta}^\dagger_k-\boldsymbol{\mu}_\theta)(\boldsymbol{\theta}^\dagger_k-\boldsymbol{\mu}_\theta)'\right\}. \tag{48}$$

Step 9. $\boldsymbol{\mu}_\beta$ を以下の条件付き分布から発生させる。

$$\boldsymbol{\mu}_\beta\,|\,\{\boldsymbol{\beta}^\dagger_k\}^K_{k=1},\,\boldsymbol{\Sigma}_\beta\sim N(\boldsymbol{m}_{\beta,1},\,\boldsymbol{S}_{\beta,1}), \tag{49}$$

ただし

$$\boldsymbol{S}^{-1}_{\beta,1}=\boldsymbol{S}^{-1}_{\beta,0}+K\cdot\boldsymbol{\Sigma}^{-1}_\beta,\ \boldsymbol{m}_{\beta,1}=\boldsymbol{S}_{\beta,1}\left(\boldsymbol{S}^{-1}_{\beta,0}\boldsymbol{m}_{\beta,0}+\boldsymbol{\Sigma}^{-1}_\beta\sum_{k=1}^{K}\boldsymbol{\beta}^\dagger_k\right). \tag{50}$$

Step 10. $\boldsymbol{\Sigma}_\beta$ を以下の条件付き分布から発生させる。

$$\boldsymbol{\Sigma}_\beta\,|\,\{\boldsymbol{\beta}^\dagger_k\}^K_{k=1},\,\boldsymbol{\mu}_\beta\sim IW\left\{r_{\beta,0}+K,\,\boldsymbol{R}_{\beta,0}+\sum_{k=1}^{K}(\boldsymbol{\beta}^\dagger_k-\boldsymbol{\mu}_\beta)(\boldsymbol{\beta}^\dagger_k-\boldsymbol{\mu}_\beta)'\right\}. \tag{51}$$

Step 11. μ_γ を以下の条件付き分布から発生させる。

$$\mu_\gamma\,|\,\{\gamma^\dagger_k\}^K_{k=1},\,\sigma^2_\gamma\sim N(m_{\gamma,1},\,s^2_{\gamma,1}), \tag{52}$$

ただし

$$s_{\gamma,1}^{-2}=s_{\gamma,0}^{-2}+K\cdot\sigma_{\gamma}^{-2},\ \ m_{\gamma,1}=s_{\gamma,1}^{2}\left(s_{\gamma,0}^{-2}m_{\gamma,0}+\sigma_{\gamma}^{-2}\sum_{k=1}^{K}\gamma_{k}^{\dagger}\right). \tag{53}$$

Step 13. σ_{γ}^{2} を以下の条件付き分布から発生させる。

$$\sigma_{\gamma}^{2}\,|\,\{\gamma_{k}^{\dagger}\}_{k=1}^{K},\,\mu_{\gamma}\sim IG\left\{\frac{r_{\gamma,0}+K}{2},\ \frac{R_{\gamma,0}+\sum_{k=1}^{K}(\gamma_{k}^{\dagger}-\mu_{\gamma})^{2}}{2}\right\}. \tag{54}$$

Step 14. ω を以下の条件付き分布から発生させる。

$$\omega\,|\,\{\phi_{k}^{\dagger}\}_{k=1}^{K}\sim G\left\{\frac{c_{0}+K\cdot n_{0}}{2},\ \frac{d_{0}+\sum_{k=1}^{K}(\phi_{k}^{\dagger})^{-1}}{2}\right\}. \tag{55}$$

Step 15. α の発生はⅣ節と同じ。

Step 16. Step 2 に戻る。

このアルゴリズムでは、事後分布を y_i^* に関して周辺化した Omori（2007）の *benchmark Gibbs sampler B* を利用している。

Ⅵ　おわりに

ディリクレ過程を事前分布に用いたベイズ推定は様々な統計的モデルに対して応用されている。特に Burgette and Reiter（2012）では、計量経済学でしばしば用いられるプロビットモデルに対して、この事前分布を用いた推定手法が提案され、議論されている。本論文では、ディリクレ過程事前分布を用いたベイズ推定の詳細について説明し、その応用例としてサンプルセレクションモデルのベイズ推定について説明した。今回は推定手法の説明に留まっているが、実際のデータを用いた分析が今後の課題となろう。

（筆者は関西学院大学経済学部専任講師）

謝辞

本研究の一部は文部科学省科学研究費（若手研究（B）課題番号 : 24730195）の助成を受けた。ここに記して、感謝の意を表する。

参考文献

Amemiya, T.（1984）. Tobit models, a survey. *Journal of Econometrics 24*（1-2）, 3-61.

Blackwell, B and J. B. MacQueen (1973). Ferguson distributions via Pólya urn schemes. *Annals of Statistics 1*(2), 353-355.

Burgette, L. F. and J. P. Reiter (2012). Multiple-shrinkage multinomial probit models with applications to simulating geographies in public use data. *Bayesian Analysis 8*(2), 453-478.

Connor, R. J. and J. E. Mosimann (1969). Concepts of independence for proportions with a generalization of the Dirichlet distribution. *Journal of the American Statistical Association 64*(325), 194-206.

Ferguson, T. S. (1973). A Bayesian analysis of some nonparametric problems. *Annals of Statistics 1*(2), 209-230.

Ghosal, S. (2001). The Dirichlet process, related priors and posterior asymptotics. In N. L. Hjort, C. Holmes, P. Müller, and S. G. Walker (Eds.), *Bayesian Nonparametrics*, Chapter 2, pp. 35-79. Cambridge: Cambridge University Press.

Ishwaran, H. and L. F. James (2001). Gibbs sampling methods for stick-breaking priors. *Journal of the American Statistical Association 96*(453), 161-173.

Ishwaran, H. and M. Zarepour (2000). Markov chain Monte Carlo in approximate Dirichlet and beta two-parameter process hierarchical models. *Biometrika 87*(2), 371-390.

Omori, Y. (2007). Efficient Gibbs sampler for Bayesian analysis of a sample selection model. *Statistics & Probability Letters 77*, 1300-1311.

Sethuraman, J. (1994). A constructive definition of Dirichlet priors. *Statistica Sinica 4*, 639-650.

安道知寛（2010）『ベイズ統計モデリング』朝倉書店。

古澄英男（2005）「ディリクレ過程事前分布を用いた構造変化のベイズ分析」和合肇編著『ベイズ計量経済分析：マルコフ連鎖モンテカルロ法とその応用』第 8 章、235-258頁、東洋経済新報社。

中妻照雄（2013）『実践ベイズ統計学』朝倉書店。

松原望（2008）『入門ベイズ統計―意思決定の理論と発展』東京図書。

和合肇編著（2005）『ベイズ計量経済分析：マルコフ連鎖モンテカルロ法とその応用』東洋経済新報社。